Recap: cut off phenomenon

1. RW on $S = \{0, 1\}^d$ very large

   $\text{(prob } \frac{1}{d+1} \text{ to stay or to flip one bit)}$

$$\textbf{Bam} \text{ ds } \frac{1}{2} \exp(-2n) \leq \| P_0^n - \mu \|_{TV} \leq \frac{1}{2\sqrt{u_0}} \cdot \exp(-\frac{2n}{d+1})$$

**Fineer analysis**

$$\| P_0^n - \mu \|_{TV} \sim 0 \quad \text{if} \quad n \gg \frac{d+1}{\log d} \quad \text{UB}$$
Proof idea:

Upper Bound (UB):
\[ \| P_0^n - \pi \|_{TV} \leq \frac{1}{2} \left( \sum_{i=0}^{2^n-1} \frac{1}{2^i} \right)^{1/2} \]
\[ \sim 0 \quad \text{as} \quad n \gg \frac{d \log d}{\epsilon} \]

Lower Bound (LB):
\[ \| P_0^n - \pi \|_{TV} = \max_{A \in S} \left| P_0^n(A) - \pi(A) \right| \]
\[ \geq \left| P_0^n(A) - \pi(A) \right| \quad \forall A \in S \]

Pick \( A_\beta = \exists \, x \in S: \quad 1 \leq \frac{d}{2} - \frac{d}{2} \leq \frac{\epsilon}{2} \sqrt{d} \)

\[ \Rightarrow \pi(A_\beta) \approx 1 \quad P_0^n(A_\beta) \approx 0 \quad \text{if} \quad n < \frac{d \log d}{\epsilon} \]
Card shuffling

$S = \{ \text{permutations of 1..N} \}$

Shuffling method = Markov chain on $S$

Question: For a given method, how long does it take to decently shuffle the deck? (i.e. to have a distribution $\varepsilon$-close in TV-dist to the uniform distribution on $S$)
Method 0: cut repeatedly the deck
not an ergodic chain (51! equivalence classes)

Method 1: "random to top"
- choose a number uniformly at random \( 1 \leq \text{rand} \leq \text{N} \)
- look for the card with this number & put it on top

This is an ergodic chain!

Claim: \( \Theta(N \log N) \) shuffles are needed with this method.
"Pf: UB"

**Coupling**

\[ X_n = \text{chain starting from the identity state } I_d \]

\[ Y_n = \text{chain starting from } T \sim \text{uniform} \]

Choose a number \( c \in \{1, N\} \) uniformly at random

& look after the card with this number in each deck, and put it on top in each deck.
Former lemma:
\[
\sup_{\pi} \left| \mathbb{P}^{n}_{\text{Id}} - \prod_{i=1}^{n} \mathbb{P}^{\pi_{i}} \right|_{TV} \leq \mathbb{P}(X_{n} \neq Y_{n})
\]

\overset{\text{dist of } X_{n}}{\text{dist of } Y_{n}}

Observation:
After each card number has been picked at least once, the two decks are the same.

So if \( T = \inf \{ n \geq 1 : \text{each number has been picked once} \} \)

Then \( \mathbb{P}(X_{n} \neq Y_{n}) \leq \mathbb{P}(T > n) \).
Caipan collector pb:

\[ E(T) = \frac{N}{N} + \frac{N}{N-1} + \frac{N}{N-2} + \cdots + \frac{N}{3} + \frac{N}{2} + \frac{N}{1} \]

\[ = N \cdot \sum_{k=7}^{N} \frac{1}{k} \approx N \cdot \log N \]

\[ \text{Var} (T) = \Theta(N^2) \]

\[ T \approx N \log N \pm N \]

So \( P(T > n) \approx 0 \) for \( n \gg N \log N \)
"\( Pf \) (LB):

\[
\sup_{\text{ACS}} \max \left| P^n_{\text{Id}}(A) - \pi(A) \right|
\geq \left| P^n_{\text{Id}}(A) - \pi(A) \right| \quad \forall \text{ACS}
\]

Choose \( A_k = \{ k \text{ bottom cards of the deck are ordered} \} \)

\[
\pi(A_k) = \frac{1}{k!} \sim \text{small}
\]

To check: \( P^n_{\text{Id}}(A) \sim 1 \) if \( n \ll N \log N \)
Observation: While $k$ cards have never been picked, at least $k$ bottom cards of the deck will be ordered (because we started from the Id permutation).

The average to pick $N-k$ different cards is

$$E(T_k) = \frac{N}{N} + \frac{N}{N-1} + \frac{N}{N-2} + \ldots + \frac{N}{k} = N \sum_{e=k}^{N} \frac{1}{e} \approx N \left( \log N - \log k \right)$$

$$\text{Var} (T_k) = \Theta(N^2)$$
Conclusion:

If \( n < < N \log N \), \( P_{\text{Id}}^{n} (A_e) \approx 1 \)

So \( \lim P_{\text{Id}}^{n} - \frac{1}{n} \mathcal{H}_{TV} \approx 1 \)

Method 2: **riffle shuffle**

\[
\Rightarrow \Theta( \log N ) \text{ Sufficient!}
\]

réf: P. Diaconis: \( \begin{array}{c}
\cdot \\
\cdot \\
\cdot \\
\cdot \\
\cdot \\
\cdot \\
\end{array} \quad (N=52) 
\]

\( \begin{array}{c}
\cdot \\
\cdot \\
\cdot \\
\cdot \\
\cdot \\
\cdot \\
\end{array} \quad n
\)
Given a distribution \((\Pi_i, i \in S)\) on a state space \(S\), how can we sample from it?

"Easy solution": let \(X\) be a r.v. with values in \(S\) such that \(P(X=i) = \Pi_i\), \(i \in S\) and assume \(S = \mathbb{N}\)
Generate $U \sim U[0,1]$ and declare

$$X = \begin{cases} 
0 & \text{if } 0 \leq U < \bar{u}_0 \\
1 & \text{if } \bar{u}_0 \leq U < \bar{u}_0 + \bar{u}_1 \\
2 & \text{if } \bar{u}_0 + \bar{u}_1 \leq U \leq \bar{u}_0 + \bar{u}_1 + \bar{u}_2 \\
\vdots 
\end{cases}$$

$$= F^{-1}(U) \quad \text{where} \quad F = \text{cdf of } X$$
Why to sample? 1. Optimization of a complex fn

\[ f : \mathbb{I}^{d} \rightarrow \mathbb{R} \]

Objective: to maximize \( f \), i.e.

to find \( x_0 \in [0,1]^d \) s.t. \( f(x_0) = \max \)

Define \( \pi(x) = \begin{cases} 0 & \text{if } f(x) \neq \max \\ c & \text{if } f(x) = \max \end{cases} \)

where \( Z = \text{number of maxima of } f \)

First idea: sample from \( \pi \) to get a maximum \( x_0 \)
Second idea: instead of sampling from $\nu_j$

Sample from $\Pi_\beta$, defined as follows:

$$\Pi_\beta(x) = \frac{e^{\beta f(x)}}{Z_\beta} \quad x \in [0,1]^d$$

where

$$Z_\beta = \sum_{x \in [0,1]^d} e^{\beta f(x)}$$

(normalization constant) \(\text{partition function}\)
2. Compute averages (Monte Carlo method)

- $X = \xi_i$, with values in $S$ and distribution $\pi_i$
  
  $\therefore \ P(X = i) = \pi_i$, $i \in S$

- $f : S \rightarrow \mathbb{R}$

- Aim: Compute $\mathbb{E}(f(x)) = \sum_{i \in S} f(i) \pi_i$

**MC method**: draw independent samples $X_1, \ldots, X_M$

- & compute $\frac{1}{M} \sum_{j=1}^{M} f(X_j) \sim \mathbb{E}(f(x))$
\[
\text{Var} \left( \frac{1}{M} \sum_{j=1}^{M} f(x_j) \right) = \frac{1}{M^2} \sum_{j=1}^{M} \text{Var} (f(x_j)) = \text{Var} (f(x))
\]

\[
\text{std dev} \sim \Theta \left( \frac{1}{\sqrt{M}} \right)
\]